Part 2: Tooth Growth Statistical Analysis

Austin L. Bistline

December 30, 2017

Overview:

In this section I perform basic inferential analysis of tooth growth, including some basic exploratory data analyses and a summary of the data. I then present confidence intervals and perform basic hypothesis testing to compare tooth growth by supp and dose. The null hypothesis is that both supplements have the same median growth outcome with each dose.

Exploring the "ToothGrowth" dataset

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.3.3
```

```
rm(list = ls()) # Reset variables to prevent errors
#?ToothGrowth # Web Summary

tg = ToothGrowth

table(tg$supp, tg$dose)
```

```
##
##      0.5  1  2
##   OJ  10 10 10
##   VC  10 10 10
```

The "tg" data set contains 60 rows of tooth-growth data corresponding to the length of odontoblasts (cells responsible for tooth growth) in 60 guinea pigs. Vitamin C was delivered to the pigs through either Orange Juice (variable "OJ") or ascorbic acid (variable "VC") in doses of 0.5, 1, or 2 mg/day. We see that six different experimental configurations were carried out, with 10 unique pigs per experiment. Many other factors, such as genes, could also be responsible for odontoblast performance, but we will assume that these other factors did not play a role.

```
summary(tg)
```

```
##       len        supp         dose
##  Min.   : 4.20   OJ:30   Min.   :0.500
##  1st Qu.:13.07   VC:30   1st Qu.:0.500
##  Median :19.25           Median :1.000
##  Mean   :18.81           Mean   :1.167
##  3rd Qu.:25.27           3rd Qu.:2.000
##  Max.   :33.90           Max.   :2.000
```

```
str(tg, tg$dose)
```

```
## 'data.frame':    60 obs. of  3 variables:
##  $ len : num  4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
##  $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
##  $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

The minimum growth length observed was 4.21, maximum was 33.9, and average was 18.81 units. Columns "len" and "dose" are numeric, and "supp" is a factor.

```
# Average tooth growth when given orange juice
mean(tg$len[tg$supp == "OJ"])
```

```
## [1] 20.66333
```

```
# Averge tooth growth when given ascorbic acid
mean(tg$len[tg$supp == "VC"])
```

```
## [1] 16.96333
```

At first glance, the teeth of guinea pigs that were given OJ grew longer on average. The next task is to determine the variance of the dataset as a function of supplement type and dose size. The goal will be to determine which supplement and/or dose is most effective for growth that is both fast and consistent. In Figure 1, box and whisker plots show at a glance the growth rate and variance of each supplement in the 3 doses that were given.

```
Dose = as.factor(tg$dose)
g = ggplot(tg, aes(x = Dose, y = len)) +
      geom_boxplot(notch = FALSE, aes(fill = Dose)) +
      facet_grid(. ~ supp) +
      scale_x_discrete("Dose (mg)") +
      scale_y_continuous("Measured Length of Odontoblasts (Teeth)")
g
```

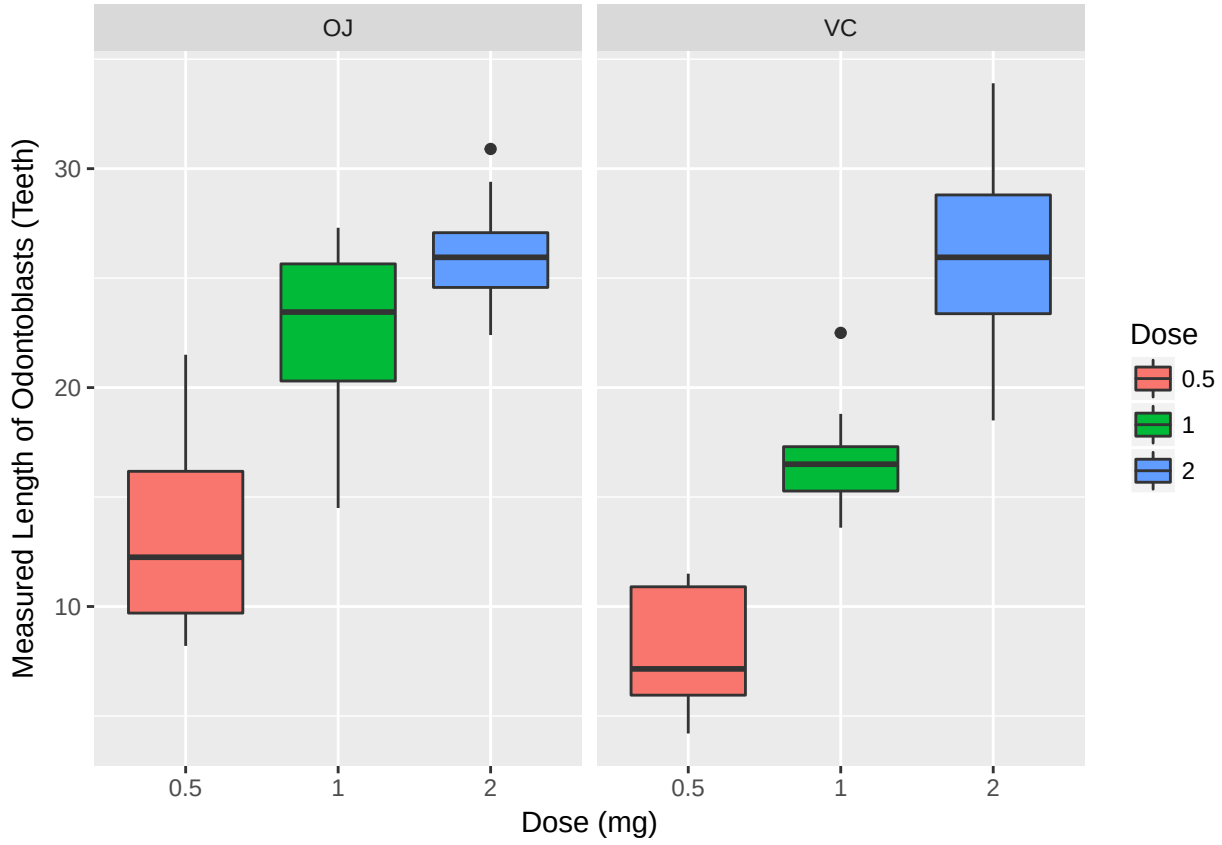

Figure 1: Guniea pig tooth-growth results given three doses (0.5, 1, and 2 mg) of two Vitamin C supplements (Orange Juice and Ascorbic Acid).

Figure 1 indicates that overall, larger doses correlate with a more growth. When OJ is given, the growth appears to be non-linear while the ascorbic acid (VC) results appear linear when examining the medians. Small and medium doses of OJ appear to be more effective for growth, while both supplements yield almost the same median growth at the largest dose size. However, the large dose ascorbic acid supplement results are much more variant. The figure indicates that if consistent maximum growth is desirable, a 2 mg/day dose of OJ will yield the best results. If consistent average growth is desirable, a 1 mg/day dose of ascorbic acid will yield the best results.

Testing the Null Hypothesis

The hypothesis testing is done with an alpha value of 0.05 ($\alpha = 0.05$) and the null will be that the two different supplements deliver the same results.

Test 1: Null implies OJ = VC @ 0.5 mg/day

```
ht1 = t.test(len ~ supp, data = subset(tg, dose == 0.5))
ht1$conf.int
```

```
## [1] 1.719057 8.780943
## attr(,"conf.level")
## [1] 0.95
```

```
ht1$p.value
```

```
## [1] 0.006358607
```

We see that the confidence interval does not include 0 and the p-value is less than α, therefore we reject the null hypothesis and accept the alternative hypothesis that OJ leads to more growth with a dose of 0.5 mg/day.

Test 2: Null implies OJ = VC @ 1 mg/day

```
ht2 = t.test(len ~ supp, data = subset(tg, dose == 1))
ht2$conf.int
```

```
## [1] 2.802148 9.057852
## attr(,"conf.level")
## [1] 0.95
```

```
ht2$p.value
```

```
## [1] 0.001038376
```

We see that the confidence interval does not include 0 and the p-value is less than α, therefore we reject the null hypothesis and accept the alternative hypothesis that OJ leads to more growth with a dose of 1 mg/day.

Test 3: Null implies OJ = VC @ 2 mg/day

```
ht3 = t.test(len ~ supp, data = subset(tg, dose == 2))
ht3$conf.int
```

```
## [1] -3.79807  3.63807
## attr(,"conf.level")
```

```
## [1] 0.95
```

```
ht3$p.value
```

```
## [1] 0.9638516
```

We see that the confidence interval includes 0 and the p-value is greater than α, therefore we cannot reject the null hypothesis that OJ and VC doses of 2 mg/day yields the same result.

Conclusion

In low doses of 0.5 to 1 mg/day, the Orange Juice supplement is more effective for tooth growth than Ascorbic Acid. At higher doses of 2 mg/day, the two supplements might be equally effective although the Ascorbic Acid results are more variable and therefore less certain.